Australian Bureau of Statistics

1351.0.55.046 - Research Paper: Assessing the Quality of Different Data Linking Methodologies Across Time, Using Northern Territory Government School Enrolment Data, Mar 2014

Summary

About this Release

This feasibility study assesses the suitability and quality of four different deterministic linkage methods, all of which link data without requiring full name and address.
Of the four linkage methods, both the Statistical Linkage Key (SLK) and the SLK+geography (SLK+) linkage methods provided a very high match-link rate and link accuracy. SLK+ is the most likely method to be implemented for linking in future data integration projects, as it has an additional linking variable that slightly improves the accuracy of the resultant dataset. A third linking method, the SLK+(edited) method, demonstrates the highly accurate linkage that can be achieved with a good quality SLK. It also highlighted that the quality of the data used to construct the linkage key directly influences the success of data linkage.
The study undertaken for students that repeated a grade shows that all four linkage methods allow the same conclusions to be drawn for the analysis questions. While there were small differences between the results achieved for each method, these differences were very small and only noticeable at the more detailed disaggregations.
This paper also details a number of considerations for improving the quality of education and training data integration projects.